JESS at thirty: strengths, weaknesses and future needs in the modelling of chemical speciation

Peter M. May

PII: S0883-2927(14)00294-7
DOI: http://dx.doi.org/10.1016/j.apgeochem.2014.11.023
Reference: AG 3376

To appear in: *Applied Geochemistry*

Please cite this article as: May, P.M., JESS at thirty: strengths, weaknesses and future needs in the modelling of chemical speciation, *Applied Geochemistry* (2014), doi: http://dx.doi.org/10.1016/j.apgeochem.2014.11.023

JESS at thirty: strengths, weaknesses and future needs in the modelling of chemical speciation

Peter M. May*

Chemistry, School of Engineering and Information Technology, Murdoch University, Murdoch, WA 6150, Australia

* Corresponding author. Tel: +61 8 9360 6068. Fax: +61 8 9360 6498

Email address: P.May@murdoch.edu.au

ABSTRACT

The current status of the software package JESS (Joint Expert Speciation System), which has been developed over the last 30 years, is described. Chemical speciation models of seawater and of metal-ion complexation in human blood plasma are used as large equilibrium systems to explore the present capabilities of the code and database. Strengths of JESS are considered to be (a) the power and flexibility of its command-driven programs, (b) the size, generality and openness of its reaction database, (c) its automatic facility to achieve thermodynamic consistency, and (d) its ability to partition chemical reactions to model kinetic constraints. A special feature of JESS is its ability automatically to generate a complete, stand-alone FORTRAN program for any particular chemical equilibrium model that has been developed. Weaknesses of JESS include the lack of a graphical user interface, the resulting effort required in familiarisation, and certain limitations regarding pressure and temperature corrections in concentrated solutions. However, the most troublesome issues - common to all 'ion-association' frameworks - are due to inadequacies in the thermodynamic data available from the literature and to persistent deficiencies in the fundamental theory of concentrated electrolyte solutions. Accordingly, work on JESS has commenced to construct a global parameterisation facility using both reaction data and the physicochemical properties of strong electrolytes in aqueous solution (including solubilities) intended to improve model function testing and to provide even better mechanisms for data harmonisation.

Keywords: chemical speciation; trace element metabolism; blood plasma; seawater; data uncertainty; weak ionic interactions

INTRODUCTION

The JESS (Joint Expert Speciation System) computer package seeks to predict the thermodynamic properties of aqueous solutions and of the phases in equilibrium with them. This paper describes the current stage of development of JESS in the context of chemical speciation modelling. It deals in particular with general equilibrium calculations by JESS based on chemical reaction thermodynamics rather than its other features which are concerned with the physicochemical properties of strong electrolytes (May et al, 2010; Rowland and May, 2010; May et al., 2011; Rowland and May, 2012; Rowland and May, 2013). Two large chemical speciation calculations using JESS (Version 8.3) will be described, one of synthetic seawater and the other of the low-molecular-weight components in human blood plasma. These systems provide rich fields for investigation, each with their own interesting lessons for speciation modellers, as well as demonstrating the computational capabilities of JESS and the extensiveness of its reaction database.

Work on JESS commenced in 1985 and many details of how the JESS software package approaches chemical speciation modelling have been described in the open literature. These include the founding raison d'etre (May and Murray, 1991a), the construction of the thermodynamic reaction database (May and Murray, 1991b), the modelling of reaction equilibria to high ionic strength and temperature (May, 2000), the attainment of thermodynamic consistency (May and Murray, 2001), a detailed case study (Filella and May, 2003), and a number of positive independent assessments (Popov and Wanner, 2005; Darn et al., 2006; Rodgers et al., 2006; Weber et al., 2006; Rodgers et al., 2007; Pak et al., 2009). However, the software package (particularly the database of chemical reactions) has evolved considerably over the years and it is apposite to take stock of its strengths and weaknesses as they now stand.

JESS is a large software package by almost any measure. It now comprises nearly 400 stand-alone programs based on some 3,000 subroutines and over 300,000 lines of FORTRAN code. The main database of thermodynamic parameters for chemical reactions (almost exclusively in aqueous solution) covers 76,500 reactions involving nearly 70,000 chemical species, together with more than 200,000 equilibrium constants, 30,000 reaction enthalpy values, 9,500 reaction Gibbs energies, 3,300 standard electrode potentials and around 2,000 reaction heat capacities. The literature database currently spans over 27,000 references. Version 8.3 also has databases with over 400,000 aqueous strong electrolyte physicochemical property values (such as activity coefficients) and nearly 50,000 solubilities.

From the outset, the focus of research using JESS was not merely to calculate chemical speciation distributions, which facility was already offered by many codes including ESTA (Equilibrium Simulation and Titration Analysis) (May et al., 1985; May et al., 1988). Rather, the goal was to grapple with a number of entrenched problems that bedevilled chemical speciation models at that time and were evidenced most vividly in the gross disparity of results obtained by groups of authoritative modellers assembled for the CHEMVAL project (Read and Broyd, 1988). Pervading every chemical speciation calculation is the issue of errors and how best to control them (May and Murray, 2001).

It seemed clear from the start that only a computational approach based on the strategy of expert systems (a technique of artificial intelligence in which specialist expertise is embedded in a decision-making process) could deal effectively with these problems. It was likewise evident that an enormous effort would be required, not just to write the computer codes but also to build the chemical databases, which could only be done by a team of researchers collaborating over many years. The chemical focus was on achieving a joint perspective between the numerous solution chemists involved: by having the databases freely open to scrutiny and correction, by having automatic methods to achieve and maintain thermodynamic consistency, by having as much flexibility as possible in the way the calculations were performed and by having powerful output facilities, e.g. to plot Pourbaix (pH - Eh) diagrams. Facile tracking of how calculated results come about was seen to be essential, regarding both assumptions made in the simulation process and links to the source input data as obtained from the chemical literature. The programming focus was on code sustainability, ensuring through modularity and other good computing practices that many programmers could contribute, that the package could be ported to inherently different types of computer hardware and operating system, and that the written code would continue to function for a long period ahead.

The main difference between JESS and other speciation modelling codes lies in the way JESS handles thermodynamic data. Instead of having an independent set of thermodynamic parameters for the formation of each 'complex' (or 'non-basis species'), the JESS reaction database code (called JTH) is designed to represent and store chemical reactions in all of their numerous forms. Assorted reaction thermodynamic parameters such as equilibrium constants, Gibbs energies, electrochemical potentials, reaction enthalpies, reaction entropies and reaction heat capacities can thus be stored *with little or no transformation by the database compilers*. This includes energies of formation as well as step-wise and overall equilibrium constants. In other words, rather than having to conform to a pre-determined format and a pre-defined set of component / master / basis species, literature data can be inserted in ways that closely match the original source.

Accordingly, instead of starting with a database that has been made thermodynamically consistent manually by (human) experts in advance, JESS inputs data that have been individually assessed (see May et al., 2011, p. 5071) but not further refined and it generates thermodynamic consistency by wholly automatic means. As illustrated in Figure 1, the calculations are performed by a sequence of programs which together comprise the so-called 'GEM Stages' (the name standing for Generalised Equilibrium Modelling). The advantages of this approach are not merely that the database contents are easier to assemble and to check but also that the database can be built and refined in a progressive and maintainable manner. This is unlike manually-constructed thermodynamic databases which, to stay consistent, must be laboriously re-processed (sometimes from scratch) when new data are entered or when corrections become necessary.

In essence, the GEM Stages are a rule-based methodology for selecting the species, the reactions and the thermodynamic parameters which best suit the user's application. Defaults are provided to make the process efficient and to guide inexperienced users but these default options can be easily altered if required. Similarly, the mass balance equations are initially formulated automatically but users can control the outcome (within mathematical constraints) to exploit different basis sets and different linear combinations of chemical reactions. However, the selection criteria encoded within

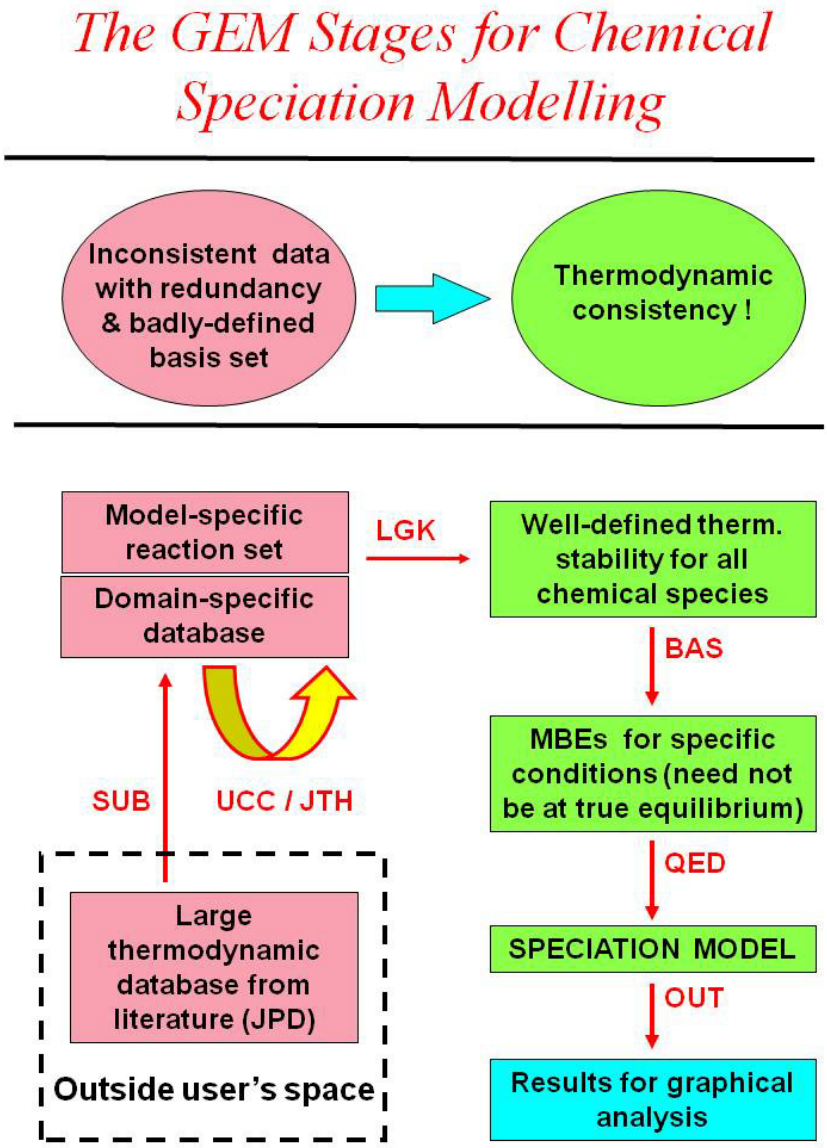

Figure 1. Schematic of the JESS chemical speciation modelling procedure, comprising the separate stages called (a) SUB, to extract reaction data for selected chemical subsystems from JPD - the 'JTH Parent Database', (b) LGK, to determine sets of appropriate basis species and to evaluate the equilibrium constants for any possible chemical reaction in the modelling domain, (c) BAS, to construct a thermodynamically consistent set of mass balance equations for a particular variable basis set, (d) QED, to solve these equations to calculate the required concentrations of all chemical species under consideration and (e) OUT, to manage various model outputs, including large scans of Eh and pH. The UCC (Unconditional Correction Coefficient) routines evaluate reaction parameters in JTH reaction databases as functions of temperature and ionic strength following Anderson et al. (1991) as described by May (2000).

JESS to find the optimum arrangement generally work well enough, particularly for new users and for experts at the start of a modelling project.

A wide variety of results relevant to modelling any aqueous solution can be obtained during and after completion of the GEM Stages calculations. For the most part, output is controlled interactively. At every Stage, a list of 'Methods and Assumptions' is generated so this can be saved as a record of what was done. A description of other output options, together with brief explanations of certain technical matters underpinning how the GEM Stages work, is given in the Appendix.

JESS MODEL OF LOW-MOLECULAR-WEIGHT COMPLEX DISTRIBUTIONS IN BLOOD PLASMA

One early investigation of a complicated multicomponent equilibrium system in aqueous solution took the form of a computer simulation of metal-ligand equilibria in blood plasma (Perrin, 1965; Hallman et al., 1971), subsequently extended to thousands of equilibria (May et al., 1977). With implications regarding the binding of metal ions to plasma proteins (May, 1990), the transport of essential elements through biological membranes (Darwish et al., 1984; May, 1990) and the way that clinical chelating therapeuticals worked *in vivo* (May and Bulman, 1983), these studies gave much impetus to experimental investigations of relevant chemical systems. Consequently, the calculated species distributions kept changing over a sustained period (e.g. Berthon et al., 1978; Blais and Berthon, 1982; Berthon et al., 1986; Brumas et al., 1993). Moreover, additional trace elements were progressively added to the original set of components (e.g. Cole et al., 1985; Dayde et al., 2003). Together with these assorted input modifications, the advent of JESS brought several possible computational extensions to the original model, most notably in an ability to deal better with redox and solubility equilibria. So, it is instructive to see how these many changes have manifested themselves over what is now an extended period of time.

Methodology

A chemical speciation model of low-molecular-weight complex distributions in blood plasma was developed using the standard facilities of the JESS software package (Version 8.3). A total of 10,818 reactions in the JESS thermodynamic database ('JPD') were identified through the SUB Stage as relevant to blood plasma with the specified composition. These reactions involved 8,938 chemical species, most of which were aqueous but included 281 solid phases. It was found (LGK Stage) that these reaction and species sets could be represented by 55 'thermodynamic basis species'. More than 50,000 equilibrium constants (or their equivalent) were found to be in scope.

Following the approach described in the original study (May et al., 1977), a well defined blood plasma composition was determined, comprising a set of total concentrations for the major components, mostly low-molecular-weight ligands (Table 1) and a set of free concentrations for the metal ions (Table 2). Holding constant the metal ion free concentrations emulates the buffering of these elements due to binding by a vast excess of plasma proteins (May et al., 1977). Other

specified conditions were the temperature = 25 °C, the pH (–log[H^+]) = 7.4, the ionic strength (*I* = 0.15 M) and the standard redox potential (Eh) = -83 mV, estimated from known concentration ratios in blood plasma of redox pairs such as cysteine / cystine and ascorbate / dehydroascorbate.

Results

The relevant data were extracted from 'JPD', the main JESS thermodynamic database for chemical reactions. In this process, any reactions present in the JPD database that are not thermodynamically consistent with those used for the speciation modelling are identified when the thermodynamic basis set and corresponding equilibrium constants are evaluated. This is done by the so-called 'Rogues Gallery' of the 'LGK' stage (see Appendix). Thus, with certain limitations discussed below, modellers are able to check much of the relevant data in a rigorous and systematic manner.

Solving the mass balance equations (in 47 Newton-Raphson iterations) was found to take about 1 second on a Dual-core laptop computer running under Windows7. The results are shown in Tables 3 and 4.

To the author's knowledge, Table 4 gives estimates of solubilities from a general thermodynamic calculation of low-molecular-weight equilibria in human blood plasma for the first time. This capability is of relevance to various possible applications of geochemistry in medicine including urolithiasis, osteoporosis and investigation of aluminium in the formation of plague associated with Alzheimer's disease (Orvig and Berthon, 1995). It is of particular interest in the context of kidney stone formation (Pak et al., 2009) where it is clinically important to calculate the saturation of certain minerals in urine.

Table 1. Total component concentrations used in the blood plasma model.

JESS Species symbol	Concentration / M	JESS Species symbol	Concentration / M
Cl-1	0.150	Histamine	3.00E-8
Na+1	0.140	Hyp-1	7.00E-6
K+1	0.004	Ile-1	6.50E-5
CO3-2	2.45E-2	Lactic-1	1.82E-3
SCN-1	1.40E-5	Leu-1	1.24E-4
NH3	2.40E-5	Lys-1	1.78E-4
PO4-3	3.81E-4	Malic-2	3.50E-5
SiH2O4-2	1.38E-4	Met-1	2.90E-5
SO4-2	2.11E-4	Orn-1	5.80E-5
2AmButan-1	2.40E-5	Oxalic-2	1.20E-5
Ala-1	3.70E-4	Phe-1	6.40E-5
Arg	9.50E-5	Pro-1	2.11E-4

Asn-1	5.50E-5	Pyruvic-1	9.50E-5
Asp-2	5.00E-6	Salicylic-2	5.00E-6
Cis-2	4.00E-5	Ser-1	1.22E-4
Citric-3	1.13E-4	Succinic-2	4.20E-5
Citrul-1	2.70E-5	Thr-1	1.50E-4
Cys-2	2.30E-5	Trp-1	1.00E-5
Gln-1	5.21E-4	Tyr-2	5.80E-5
Glu-2	4.80E-5	Val-1	2.27E-4
Gly-1	2.43E-4	H+1_Ascorbic-2	4.30E-5
His-1	8.50E-5		

Table 2. Free fixed metal ion concentrations used in the blood plasma model.

Metal ion	Concentration / M
Ca^{2+}	0.00114
Cd^{2+}	1.E-15
Cu^{2+}	1.E-18
Fe^{2+}	1.E-23
Pb^{2+}	1.E-14
Mg^{2+}	0.00052
Mn^{2+}	1.E-12
Ni^{2+}	1.E-17
Zn^{2+}	1.E-09

Table 3. Percentage distributions at t = 25 °C and pH = 7.4 of metal ions amongst low-molecular weight ligands in human blood plasma as found in the 1977 study (May et al., 1977) and in the present work.

Metal		**Percentage**[a,b]	
Complex (chemical name[c] and JESS symbol)	**Charge**	**1977 study[d]**	**This work[e]**
Cadmium - Cd(II)			
Cysteinate - Cd+2_Cys-2	0	-	98
Bis(cysteinate) - Cd+2_Cys-2(2)	–2	-	1
Calcium - Ca(II)			

Protonated carbonate - Ca+2_H+1_CO3-2	+1	9	7
Citrate - Ca+2_Citric-3	–1	4	3
Lactate - Ca+2_Lactic-1	+1	3	2
Phosphate - Ca+2_PO4-3	–1	3	0
Carbonate - Ca+2_CO3-2	0	2	1
Chloride - Ca+2_Cl-1	+1	-	12
Copper - Cu(I)[f]			
Cysteinate - Cu+1_Cys-2	–1	-	77
Cysteinate chloride - Cu+1_Cl-1_Cys-2	–2	-	15
Protonated cysteinate - Cu+1_H+1_Cys-2	0	-	5
Thiocyanate - Cu+1_SCN-1	0	-	3
Copper - Cu(II)			
Cystinate histidinate - Cu+2_His-1_Cis-2	–1	21	2
Protonated cystinate histidinate - Cu+2_H+1_His-1_Cis-2	0	17	1
Bis(histidinate) - Cu+2_His-1(2)	0	11	14
Histidinate threoninate - Cu+2_His-1_Thr-1	0	8	9
Histidinate valinate - Cu+2_His-1_Val-1	0	5	4
Protonated histidinate lysinate - Cu+2_H+1_His-1_Lys-1	+1	5	3
Alanate histidinate - Cu+2_Ala-1_His-1	0	4	6
Histidinate serinate - Cu+2_His-1_Ser-1	0	4	4
Histidinate phenylalanate - Cu+2_His-1_Phe-1	0	3	2
Glycinate histidinate - Cu+2_Gly-1_His-1	0	3	5
Histidinate leucinate - Cu+2_His-1_Leu-1	0	2	3
Glutamate histidinate - Cu+2_His-1_Glu-2	–1	2	2
Glutaminate histidinate - Cu+2_Gln-1_His-1	0	2	25
Iron - Fe(II)[f]			
Citrate hydroxide - Fe+2_OH-1_Citric-3	–2	-	98
Carbonate - Fe+2_CO3-2	0	-	1
Iron - Fe(III)			
Citrate hydroxide	–1	99	91
Bis (citrate hydroxide) - Fe+3_H+1(-2)_Citric-3(2)	-5	-	5
Glutamate hydroxide - Fe+3_OH-1_Glu-2	0	-	1
Lead - Pb(II)			
Cysteinate - Pb+2_Cys-2	0	80	96
Citrate cysteinate - Pb+2_Citric-3_Cys-2	–3	7	0
Protonated cystinate - Pb+2_H+1_Cis-2	+1	5	1
Protonated cysteinate phosphate - Pb+2_H+1_Cys-2_PO4-3	–2	3	0
Protonated carbonate - Pb+2_H+1_CO3-2	+1	2	0

Protonated bis(cysteinate) - Pb+2_H+1_Cys-2(2)	–1	2	1
Carbonate - Pb+2_CO3-2	0	-	1
Magnesium - Mg(II)			
Protonated carbonate - Mg+2_H+1_CO3-2	+1	6	5
Citrate - Mg+2_Citric-3	–1	5	2
Carbonate - Mg+2_CO3-2	0	2	1
Lactate - Mg+2_Lactic-1	+1	2	1
Protonated phosphate - Mg+2_H+1_PO4-3	0	1	1
Chloride - Mg+2_Cl-1	+1	-	15
Manganese - Mn(II)			
Protonated carbonate - Mn+2_H+1_CO3-2	+1	24	32
Citrate - Mn+2_Citric-3	–1	10	3
Carbonate - Mn+2_CO3-2	0	2	14
Oxalate - Mn+2_Oxalic-2	0	2	0
Protonated phosphate - Mn+2_H+1_PO4-3	0	1	4
Chloride - Mn+2_Cl-1	+1	-	6
Zinc - Zn(II)			
Citrate cysteinate - Zn+2_Citric-3_Cys-2	–3	43	0
Bis(cysteinate) - Zn+2_Cys-2(2)	–2	19	29
Cysteinate histidinate - Zn+2_His-1_Cys-2	–1	12	17
Cysteinate - Zn+2_Cys-2	0	3	9
Histidinate - Zn+2_His-1	+1	3	16
Protonated bis(cysteinate) - Zn+2_H+1_Cys-2(2)	–1	1	3
Bis(histidinate) - Zn+2_His-1(2)	0	1	7
Cysteinate glutaminate - Zn+2_Gln-1_Cys-2	–1	1	2
Citrate - Zn+2_Citric-3	-1	-	3

[a] Percentage of total metal ion concentration

[b] Species accounting for less than 1% are omitted

[c] The chemical names follow the original study [c]

[d] May et al., 1977.

[e] pe = 1.40; Eh = 83 mV

[f] Metal ion not included previously [c]

Table 4. Calculated saturation indices for solids in the vicinity of supersaturation as determined by the blood plasma model[a].

Solid species (chemical name and JESS symbol)	**lg(Ksp')[b]**	**lg(SI)[c]**
Hydrated octacalcium phosphate OCP: Ca+2(4)_H+1_PO4-3(3)_H2O(3)_(s)	12.75	-0.88

Brushite: Ca+2_H+1_PO4-3_H2O(2)_(s)	0.88	-0.88
Weddellite: Ca+2_Oxalic-2_H2O(2)_(s)	-7.41	-0.81
Ikaite: Ca+2_CO3-2_H2O(6)_(s)	3.35	-0.57
Calcium oxalate trihydrate: Ca+2_Oxalic-2_H2O(3)_(s)	-7.75	-0.47
Whewellite: Ca+2_Oxalic-2_H2O_(s)	-7.83	-0.39
Ca+2(2)_Fe+2_PO4-3(2)_H2O(4)_(s)	-6.64	-0.15
Vaterite: Ca+2_CO3-2_(Vater.,s)	2.93	-0.15
Calcium citrate tetrahydrate: Ca+2(3)_Citric-3(2)_H2O(4)_(s)	-18.47	-0.07
Sodium hydrogen carbonate: H+1_Na+1_CO3-2_(s)	-2.69	0.14
Calcite: Ca+2_CO3-2_(Calc.,s)	2.40	0.38
Aragonite: Ca+2_CO3-2_(Arag.,s)	2.55	0.23

[a] Anhydrous solids known to be reluctant or unable to precipitate from aqueous solutions under ambient conditions (Königsberger et al., 1999; Gautier et al., 2014) have been omitted from the Table; these include monetite, magnesium phosphate, magnesite, dolomite (ordered and disordered), calcium oxalate, and several anhydrous calcium phosphates.

[b] Conditional solubility product in terms of the relevant basis species concentrations as $\log_{10}$ value.

[c] Solubility index (concentration product / Ksp') as $\log_{10}$ value.

JESS MODELS OF SYNTHETIC SEAWATER

Seawater models were the very first chemical speciation calculations performed for complicated multicomponent equilibrium systems. Sillén (1961) described the concept originally and subsequently used such equilibrium calculations to propose how seawater got its present composition (Sillén, 1967a; Sillén, 1967b). Garrels and Thompson (1962) developed a more detailed chemical model for seawater, with emphasis on activity coefficients. This work was extended by Whitfield (1973) who applied the Brønsted-Guggenheim specific ion-interaction framework to deal better with multicomponent mixtures. The use and validity of concentration quotients in the context of seawater modelling was then described by Dickson et al. (1981), leading to an ion-association model for electrolyte mixtures for the major components in seawater (Dickson and Whitfield, 1981) and an authoritative calculation of the chemical speciation of numerous seawater trace elements (Turner et al., 1981). A number of refinements followed (Turner and Whitfield, 1987; Byrne et al., 1988; Clegg and Whitfield, 1995). Byrne et al. (1988) examined the influence of temperature and pH on trace metal speciation.

The tradition of seawater modelling thus seems well established. It may, therefore, be assumed that a sound foundation exists in respect of the most appropriate methodology, the component concentrations to use, and the likely outcomes to be expected. Unfortunately, this is not the case.

Table 5. The (bulk) composition of seawater used in previous modelling studies[a] by Garrels and Thompson (1962), Sillén, (1967b), Dickson and Whitfield (1981), Turner et al. (1981)*, Clegg and Whitfield (1995), and Crea et al. (2006).

Component name	Symbol	Component total concentrations / M						
		1962[b]	1967	1980	1981a[c,d]	1981*	1995	2006[f]
Borate	$B(OH)_4^-$		4×10^{-4}			4.27×10^{-4}	4.4×10^{-4}	
Bromide	Br^-		8×10^{-4}		8.7×10^{-4}	8.51×10^{-4}	8.7×10^{-4}	
Calcium	Ca^{2+}	1.04×10^{-2}	1.03×10^{-2}	1.063×10^{-2}	1.065×10^{-2}	0-1.047×10^{-2}	1.064×10^{-2}	1.11×10^{-2}
Carbonate	CO_3^{2-}	2.65×10^{-3}	2.3×10^{-3}	2.285×10^{-3} [e]		19.5-0.65×10^{-3}	2.41×10^{-3}	
Chloride	Cl^-		0.5459	0.05657	0.56825	0.56234	0.56579	0.5639
Fluoride	F^-		1×10^{-4}		7×10^{-5}	6.9×10^{-5}	6×10^{-5}	
Potassium	K^+	1.00×10^{-2}	9.9×10^{-3}	1.062×10^{-2}	1.058×10^{-2}	1.047×10^{-2}	1.058×10^{-2}	1.10×10^{-2}
Magnesium	Mg^{2+}	5.40×10^{-2}	5.32×10^{-2}	5.489×10^{-2}	5.475×10^{-2}	5.37×10^{-2}	5.519×10^{-2}	5.48×10^{-2}
Sodium	Na^+	0.4752	0.4680	0.4822	0.48617	0.47863	0.48525	0.4211
Strontium	Sr^{2+}		1×10^{-4}		9×10^{-5}	8.9×10^{-5}	9×10^{-5}	
Sulfate	SO_4^{2-}	2.84×10^{-2}	2.82×10^{-2}	2.906×10^{-2}	2.927×10^{-2}	2.884×10^{-2}	2.927×10^{-2}	2.88×10^{-2}

[a] Note 1: Byrne et al. (1988) give only total concentrations ($g.kg^{-1}$) for 'ligands' ; carbonate = 2.14×10^{-3}, chloride = 0.56, fluoride = 3.5×10^{-5} and sulfate = 0.0095; sulfate is set at the free ion concentration calculated by Turner et al. (1981) - see Byrne et al. (1988) Table 1, footnote b.

Note 2: The concept of ionic strength (I) is used ambiguously in this literature.

[b] $I = 0.70$

[c] $I = 0.65$

[d] Alkalinity = 2.0 - 22.4×10^{-4} M

[e] Sum of CO_3^{2-} and HCO_3^- concentrations

[f] Synthetic seawater

Table 6. The bulk composition of seawater solutions used in the present study compared with that of 'average[a] / standard[b] seawater' as reported by Millero (2002) and Millero et al. (2008).

Component[c]	Symbol	Selected (mol./L)	Selected (mol. / kg H_2O)	2002 (mol. / kg H_2O)	2008 (mol. / kg H_2O)
Borate	$B(OH)_4^-$	4.266×10^{-4}	4.315×10^{-4}	1.0×10^{-4}	4.303×10^{-4}
Bromide	Br^-	8.511×10^{-4}	8.651×10^{-4}	8.4×10^{-4}	8.728×10^{-4}
Calcium	Ca^{2+}	1.047×10^{-2}	1.059×10^{-2}	1.028×10^{-2}	1.0657×10^{-2}
Carbonate	CO_3^{2-}	2.01×10^{-3}	2.03×10^{-3}	2.02×10^{-3}	2.0380×10^{-3}
Chloride	Cl^-	0.56234	0.56873	0.54588	0.565765
Fluoride	F^-	6.918×10^{-5}	7.00×10^{-5}	7×10^{-5}	7.08×10^{-5}
Potassium	K^+	1.047×10^{-2}	1.060×10^{-2}	1.021×10^{-2}	1.0580×10^{-2}
Magnesium	Mg^{2+}	5.370×10^{-2}	5.431×10^{-2}	5.282×10^{-2}	5.4742×10^{-2}
Sodium	Na^+	0.47863	0.48653	0.46907	0.486060
Strontium	Sr^{2+}	8.913×10^{-5}	9.01×10^{-5}	9×10^{-5}	9.40×10^{-5}
Sulfate	SO_4^{2-}	2.884×10^{-2}	2.917×10^{-2}	2.824×10^{-2}	2.9264×10^{-2}

[a] Salinity = 35, Chlorinity - 19.374, pH_{SWS} = 8.1, TA = 2.400 mmol kg^{-1} , t = 25 °C

[b] Salinity = 35, TA = 2.300, mmol kg^{-1} , CO_2(g) fugacity = 33.74 Pa = 3.33×10^{-4} atm., t = 25 °C

[c] It is interesting to note that silicon has been omitted from these definitions.

Methodology

To develop an unambiguous thermodynamic model, a well-defined set of *bulk* seawater component concentrations needs first to be specified. Following the compositions used in earlier studies (Table 5), the values shown in Table 6 were selected for use in all calculations of seawater performed here. Preference was given to (Turner et al., 1981) because this paper provided a detailed set of results that could be useful for comparisons. To this set of bulk components were added various trace elements at appropriate levels determined from the literature (Lide, 2005), as shown in Table 7. Other specified values included the temperature = 25 °C, the pH ($-\log[H^+]$) = 8.2, the formal ionic strength (I = 0.65, corresponding to a salinity of ~32 ‰) and the redox potential (Eh / mV), imposed as outlined below for the investigation of three main Scenarios (named A, B and C respectively) and without invoking the charge balance equation.

Investigations of carbonate chemistry speciation in seawater are typically based on measured DIC (dissolved inorganic carbon) and TA (total alkalinity) adopting the stoichiometric equilibrium constants for carbonic acid *in seawater* determined by Mehrbach et al. (1973) and refitted by Dickson and Millero (1987). The accuracy to which alkalinity titration data can then be fitted by an equilibrium model is discussed in excellent detail by Butler (1992). However, this has not been the approach taken here since the objective of this work is to investigate JESS as a *general* chemical speciation facility, not intended to emulate the properties of seawater by empirical fitting. Thus, the solution being modelled is a synthetic seawater uncomplicated by presence of organic substances (which are known to be significant metal-binding components in the natural oceans but are, at least in part, thermodynamically uncharacterisable). Moreover, the equilibrium constants employed by JESS are measured only in simple background electrolytes like NaCl, $NaClO_4$ and KNO_3, not in a seawater medium (real or artificial). As foreseen by Byrne et al. (1988), the subtleties in inorganic speciation schemes thereby exposed are (for the time being) of greater chemical than oceanic interest.

Scenario A considers highly oxidizing conditions, representing those approached by equilibration with atmospheric oxygen; then, at pH = 8.200, pe = +11.535 and Eh = +682 mV, corresponding to a partial pressure $p[O_2(g)] = 2.5 \times 10^{-4}$ atm or 8.0 mg/L (8 mg of gas in equilibrium above one litre of seawater solution). As a comparison, Turner et al. (1981) give a limit for 'oxic' seawater solutions as pe < +12.4. Casting the mass balance equations in terms of SO_3^{2-} as the basis species to represent the non-sulfate sulfur-containing species confers numerical robustness on the modelling calculations which is always convenient (and sometimes necessary).

Scenario B considers highly reducing conditions, representing in redox terms the mirror image of Scenario A or, in other words, the state of seawater approached by equilibration with hydrogen gas; then, at pH = 8.200, pe = –6.586 and Eh = –390 mV, corresponding to $p[H_2(g)] = 2.5 \times 10^{-4}$ atm or 0.5 mg/L. As a comparison, Turner et al. (1981) give an upper limit for 'anoxic' seawater solutions as pe > -5.0. In Scenario B the species HS^- was used to represent the other non-sulfate sulfur-containing species (rather than the choice of SO_3^{2-} as the basis species under Scenario A).

Scenario C considers an intermediate redox condition, half way between Scenarios A & B; then, at pH = 8.200, pe = +2.5 and Eh = +148 mV, corresponding to negligible concentrations of both $O_2(g)$

and H_2(g). This Scenario was then considered in two ways, labelled Scenario C1 and Scenario C2, the only difference between these being the way the mass balance equations were set up and solved mathematically, namely that HS^- and SO_3^{2-} respectively were each separately considered as basis species. Since a specific choice should not change the calculated chemical speciation, this provides a an important internal test of the model.

Scenario D considers the same conditions as Scenario C except that the pH = 8.000.

For reasons that are explained below, all of the modelling Scenarios A - D were also re-calculated under the modelling assumption that the chloride species of sodium, potassium, magnesium, calcium and strontium ions should be ignored. These additional Scenarios were labelled A' - D' respectively.

The JESS facilities for thermodynamic consistency checking were employed with all investigated modelling Scenarios as described above for blood plasma.

Results

A chemical speciation model of seawater was thus developed using the standard facilities of the JESS software package (Version 8.3). At total of 2,318 reactions in the JESS thermodynamic database ('JPD') were identified as relevant to seawater with the specified composition. These reactions involved 1,274 chemical species, 754 of which were aqueous and 486 represented solid phases. It was found that these reaction and species sets could be represented by 22 so-called 'thermodynamic basis species' (meaning that 22 independent variables are needed to describe fully the operative thermodynamic relationships between all the remaining species). More than 20,000 equilibrium constants (including Gibbs energies and standard potentials) were found to be in scope.

A modelling decision to treat sulfate, SO_4^{2-}, separately from other sulfur-containing species was made. This was implemented by imposing a kinetic constraint on the chemical reactions of sulfate, having the practical effect of increasing the number of basis species from 22 to 23. The identity of the additional sulfur-containing basis species was varied under different modelling Scenarios to suit the extreme ranges of redox potential, 'Eh', that were to be investigated. Under reducing conditions, the mass balance equation (additional to sulfate) was formulated in terms of the singly-protonated sulfide anion, HS^-, and under oxidizing conditions it was formulated in terms of sulfite, SO_3^{2-}. The separation of SO_4^{2-} from HS^-, or SO_3^{2-}, in this manner assumes simply that, in the absence of microbial catalysts, SO_4^{2-} cannot be rapidly reduced. This is in accord with practical experience. From a modelling perspective, it allows the interactions of sulfate present at relatively high concentrations in seawater to be explored independently of other less abundant sulfur-containing species, which are known to be more readily oxidized and reduced, depending on the Eh of the solution. This could be important because some of these other sulfur-containing species, especially sulfide, interact strongly with transition metal ions.

With each of the modelling Scenarios defined above, solving the mass balance equations (in ~30 Newton-Raphson iterations) was found to take about 0.2 seconds on a Dual-core laptop computer

running under Windows7. The results are shown in Table 8. Since the outcome of the calculation must be the same regardless of choice of basis species, the results shown in Tables 9 and 10 (Scenarios C1 and C2) give assurance that the kinetic constraint imposed on the above seawater models in respect of sulfur species functions properly.

Table 7. The minor components of seawater solutions used in the present study.

Component name	Symbol	Component total concentrations (mol./L)
Aluminium	Al^{3+}	7.4×10^{-8}
Cobalt	Co^{2+}	$3.4 \rangle 10^{-10}$
Copper	Cu^{2+}	$3.9 \rangle 10^{-9}$
Iron(III)	Fe^{3+}	$3.6 \rangle 10^{-8}$
Manganese	Mn^{2+}	$3.6 \rangle 10^{-9}$
Nickel	Ni^{2+}	$9.5 \rangle 10^{-9}$
Sulfur (other)[a]	SO_3^{2-}	$2.0 \rangle 10^{-4}$
Sulfur (other)[b]	HS^-	$2.0 \rangle 10^{-4}$
Silicon	$SiH_2O_4^{2-}$	$7.8 \rangle 10^{-5}$
Zinc	Zn^{2+}	$7.5 \rangle 10^{-8}$

[a] Scenarios A, A', C1, C1' and D

[b] Scenarios B, B', C2 and C2'

Table 8. Major species (as percentages of total element concentrations) calculated by different modelling Scenarios.

Element	Major Species	A	B	C1	C2	D	A'	B'	C1'	C2'	D'
Al	Al+3_OH-1(4)	76	76	76	76	72	75	75	75	75	72
B	B(OH)3	78	78	78	78	85	77	77	77	77	84
Br	Br-1	99	99	99	99	99	99	99	99	99	99
C (inorganic)	H+1_CO3-2	71	71	71	71	74	69	69	69	69	72
Ca	Ca+2	56	56	56	56	56	73	73	73	73	73
Cl	Cl-1	93	93	93	93	93	100	100	100	100	100
Co	Co+2	51	7	51	51	55	50	7	50	50	54
Cu	Cu+1_Cl-1(2)	-	-	61	61	62	-	-	55	55	56
	Cu+2_CO3-2	60	-	-	-	-	57	-	-	-	-
	Cu+1_H+1_S-2(2)	-	99	-	-	-		99	-	-	-
F	F-1	58	58	58	58	58	56	56	56	56	56
Fe	Fe+3_OH-1(3)	79	-	77	77	77	79		77	77	77
	Fe+2_H+1_S-2	-	90	-	-	-	-	90	-	-	-

K	K+1	82	82	82	82	82	94	94	94	94	94
Mg	Mg+2	81	81	81	81	82	89	89	89	89	89
Mn	Mn+2	-	45	45	45	45	43	-	43	43	43
Mn	Mn+3_OH-1(2)	98	-	-	-	-	98	-	-	-	-
Mn	Mn+2_H+1_S-2	-	83	-	-	-	-	83	-	-	-
Na	Na+1	92	92	92	92	92	99	99	99	99	99
Ni	Ni+2	56	56	56	56	58	55	-	55	55	56
	Ni+2_H+1_S-2	-	89	-	-	-	-	89	-	-	-
S in sulfate	SO4-2	51	51	51	51	51	48	48	48	48	48
S not in sulfate	H+1_S-2	-	97	-	-	-	-	97	-	-	-
	SO3-2	41	-	41	41	41	41	-	39	39	39
Si	H+1(2)_SiH2O4-2	96	96	96	96	97	95	95	95	95	97
Sr	Sr+2	70	70	70	70	70	85	85	85	85	85
Zn	Zn+2	53	-	53	53	53	52	-	52	52	53
	Zn+2_H+1_S-2(2)	-	100	-	-	-	-	100	-	-	-

Table 9. Solution of equations in Scenarios C1 (left) and C2 (right) showing identical concentrations for all basis species in common. Changed basis species concentration indicated by an asterisk, *.

Scenario C1		**Scenario C2**	
mol./L	**Species**	**mol./L**	**Species**
5.51E-05	B(OH)4-1	5.51E-05	B(OH)4-1
0.000841	Br-1	0.000841	Br-1
0.005878	Ca+2	0.005877	Ca+2
0.523429	Cl-1	0.523429	Cl-1
1.74E-10	Co+2	1.74E-10	Co+2
1.84E-13*	Cu+2	6.76E-14*	Cu+1
4.04E-05	F-1	4.04E-05	F-1
7.98E-20*	Fe+3	4.94E-10*	Fe+2
0.008559	K+1	0.008559	K+1
4.375E-02	Mg+2	4.375E-02	Mg+2
1.60E-09	Mn+2	1.60E-09	Mn+2
0.441254	Na+1	0.441254	Na+1
5.31E-09	Ni+2	5.31E-09	Ni+2
8.15E-5*	SO3-2	1.E-41*	H+1_S-2
0.01471	SO4-2	0.01471	SO4-2
6.28E-05	Sr+2	6.28E-05	Sr+2
4.01E-08	Zn+2	4.01E-08	Zn+2
6.79E-11	Al+3_OH-1(3)	6.79E-11	Al+3_OH-1(3)
7.45E-05	H+1(2)_SiH2O4-2	7.45E-05	H+1(2)_SiH2O4-2
0.001423	H+1_CO3-2	0.001423	H+1_CO3-2

Table 10. Chemical speciation profiles (percentage distributions) and species concentrations for iron and copper in Scenario C, these being independent of basis set, with identical results from models C1 and C2.

Component	Species	%[a]	Concs. / M
Cu(II)	Cu+2_CO3-2	60	3.38E-12
	Cu+2_OH-1_CO3-2	12	6.95E-13
	Cu+2_CO3-2(2)	9	5.15E-13
	Cu+2_OH-1	5	2.99E-13
	Cu+2_B(OH)3	3	1.86E-13
	Cu+2	3	1.84E-13
	Cu+2_OH-1(2)	3	1.43E-13
	Cu+2_Cl-1	2	1.25E-13
Cu(I)	Cu+1_Cl-1(2)	61	2.39E-09
	Cu+1_Cl-1(4)	20	7.91E-10
	Cu+1_Cl-1(3)	12	4.82E-10
	Cu+1_Cl-1_OH-1	3	1.06E-10
	Cu+1_Cl-1	2	6.58E-11
	Cu+1_Cl-1(2)_OH-1	1	4.62E-11
Fe(III)	Fe+3_OH-1(3)	79	2.78E-08
	Fe+3_OH-1(4)	20	7.07E-09
	Fe+3_OH-1(2)	1	1.87E-10
Fe(II)	Fe+2	55	4.94E-10
	Fe+2_CO3-2	18	1.62E-10
	Fe+2_Cl-1	14	1.28E-10
	Fe+2_Cl-1(2)	7	5.90E-11
	Fe+2_SO4-2	5	4.42E-11
	Fe+2_OH-1	1	1.33E-11

[a] Species accounting for >1% have been omitted

Remarkably few studies of the chemical speciation in seawater give tables of percentage distributions. This no doubt reflects the uncertainties embedded in seawater thermodynamics outlined below in this paper. However, it is possible to make some direct comparisons with the early studies in this area (which were, in fact, seminal in changing the focus away from such percentage distributions). The results are shown in Table 11.

Table 11. Calculated percentages of trace metals associated with particular seawater ligands (Scenario C1).

Component	Ligand	This work	Turner et al. (1981)	Byrne et al. (1988)
Al	Hydroxide	100[a]	100	100
Cu(I)	Free ion	0	0	0
	Chloride	100[b]	100	100
Cu(II)	Free ion	3	9	5
	Carbonate	81[c]	79	85
	Sulfate	0	1	0
	Hydroxide	5	8	8
	Chloride	2	3	2
Fe(II)	Free ion	55	69	69
	Sulfate	5	4	5
	Hydroxide	1	2	2
	Chloride	14	20	13
	Carbonate	18	5	10
Fe(III)	Hydroxide	100	100	100

[a] Includes a fluoride-hydroxy mixed complex (8%)
[b] Includes a chloride-hydroxy mixed complex (3%)
[c] Includes a carbonate-hydroxy mixed complex (12%)

Several other properties of practical oceanographic significance are calculated directly by JESS and are shown in Table 12. Corresponding values for seawater from the literature are also given for the purposes of comparison discussed below. Gypsum solubility is selected deliberately because it is well known to cause trouble for ion-association models of seawater and brines (Harvie and Weare, 1980).

Table 12. Model predictions for synthetic seawater solutions

Quantity	JESS result Scenario C1	JESS result Scenario C1'	Literature value[a]
Calcite sat. index	3.6	4.6	5[b]
Aragonite sat. index	2.6	3.3	3[b]
pCO_2	3.73×10^{-4} atm	3.63×10^{-4} atm	~ 4×10^{-4} atm

Alkalinity	100 mg.L^{-1} $CaCO_3$ = 2.0 mEq.L^{-1}	100 mg.L^{-1} $CaCO_3$ = 2.0 mEq.L^{-1}	2300 - 2350[b] µmol.kg^{-1}
Gypsum solubility	0.081 mol.L^{-1}	0.074 mol.L^{-1}	0.012[c,d] mol.L^{-1}

[a] Natural (surface) seawater
[b] Millero (2007)
[c] Rolnick (1954)
[d] Linke (1965)

DISCUSSION

Assessments of software packages based on comparisons of their final results using some physical measurement as a yardstick are less helpful than they may perhaps seem because 'success' may hinge on just a single value in a large database, perhaps unrepresentative of data quality as a whole or (even more problematically) because it has been evaluated for that specific purpose. Obviously, with large numbers of adjustable parameters available, almost any given experimental observation can be closely matched by a targeted numerical regression procedure. Serious difficulties thus arise in any quest to find objective tests to judge model calculations because it is difficult to distinguish 'correlations' (in which the experimental data are used to parameterise the model) and 'predictions' (in which they are not). For the purposes of this paper, therefore, strengths and weaknesses will be canvassed broadly by considering the range of results described above for the blood plasma and seawater models.

To paraphrase a recent investigation of mineral solution thermodynamics (Zittlau et al., 2013): the diversity of minerals in nature, and the variability of their behaviour under different conditions, raises difficult questions about their stability and precipitation sequences. Such questions can be addressed by thermodynamic modelling codes, a profusion of which have been described in the literature. "These programs calculate chemical speciation in aqueous solutions and evaluate the saturation state of these solutions with respect to selected solids. The limitations of thermodynamic modelling applied in natural waters arise from the completeness and accuracy of the input data (chemical composition of natural waters) as well as from the availability and reliability of the appropriate thermodynamic data for solids and aqueous species." (Zittlau et al., 2013).

The blood plasma model

It is now feasible to solve very large equilibrium systems without recourse to codes like ECCLES (May et al., 1977) which was tailor-made to deal with sparse matrices. The JESS approach to solving mass balance equations using a well-chosen set of basis variables is both efficient and robust and, in this author's opinion, is far superior to the theoretically equivalent but numerically more challenging minimization of Gibbs energies. Moreover, in contrast to the early blood plasma simulation, the vast

effort needed to assemble the thermodynamic input data for chemical speciation calculations of such dimension has now been automated. This has conferred many advantages in addition to speed and convenience: it is reasonable to believe that there are fewer compiler-introduced errors than would be the case otherwise and that the entire modelling process is more sustainable.

While there have been significant changes to the computed distributions of metal ions amongst low-molecular-weight ligands over time, these can be attributed without exception to changes in the equilibrium constants available, mostly coming from further experimentation. Thanks to this collective experimental endeavour, many of the chemical systems required to model the aqueous solution chemistry of metal-ligand equilibria in biofluids are now much better characterised. However, the available thermodynamic data for a number of apparently significant species, such as for calcium-bicarbonate, remains far from ideal.

Table 3 shows the species that were identified in early blood plasma models but are no longer considered to be predominant. Anyone imagining that the value of these calculations lay in the precise percentages initially reported would have soon been disappointed (Berthon et al., 1978). From the outset, the most pressing tasks required to improve the blood plasma model, including the need for better data, were well appreciated (May et al., 1977). On the other hand, despite these quantitative limitations, the modelling led immediately to a new way of thinking about metal ion equilibria *in vivo* which crystallised into enduring concepts (May, 1990) such as how metal ions could be exchanged between macromolecules in solution, how the low-molecular-weight complex fraction might change the flux of transition metal ions between body compartments, and how mixed ligand complexes were likely to be biologically important. None of these ideas depended on the exact percentage distributions that were calculated at that time but they have nevertheless had a broad and long-lasting impact on the development of bioinorganic medicine (Berthon, 1995, Berthon and Lambs, L., 1995, Davidge, J. and Williams, D.R., 2003, Filella, 1995, Jackson, 1995a, Jackson, 1995b).

Since sodium and chloride ions are now explicitly included in the JESS models of biofluids (cf. ECCLES), some ion pairing can be seen in the blood plasma results but, notably, this is more-or-less negligible because of the relatively low NaCl(aq) concentration (cf. seawater and hypersaline brines).

The seawater models

The results of the seawater models indicate that, even with well-studied equilibrium systems, current predictions are qualitatively but not quantitatively satisfactory. Apart from the reactions of a few redox-sensitive metal ions, seawater has a simple solution chemistry yet one which defies accurate characterisation using general (i.e. non-empirical) thermodynamic methods and ordinary (i.e. not bespoke) thermodynamic parameters.

In modelling the thermodynamics of natural systems, absolute calculations and/or estimates rarely prove to be of quantitative value, in contrast to relative values obtained from comparable calculations. This can be seen in Table 10 (where percentages of the individual species are

compared) and in Table 11 (where overall complexation by particular ligands provides the basis for comparison). However, even this portrayal doesn't help that much. It is clear that all available models agree that aluminium and iron(III) speciation at pH = 8.2 under oxidizing conditions will be entirely dominated by hydroxy-complexes but such a 'prediction' is merely in accord with what would be considered obvious by most co-ordination chemists. The results for other elements are much less obvious and this is fully reflected in the spread of model calculations. It can therefore be concluded that, as with the blood plasma model, there are often poor alignments between the percentage distributions of species calculated over time and by different software packages. Even with the predominant species, where some agreement is evident, the calculated percentages can vary widely. Reasons for these fluctuations in chemical speciation calculations are considered in some detail below.

Solubility predictions

There is a very well established empirical rule for predicting sequences of precipitates formed from solutions that deserves greater attention than it often gets. Called 'Ostwald's rule of stages' (Ostwald, 1897), it holds that the least thermodynamically stable solid will appear first from a supersaturated solution. Although this may seem counter-intuitive, it is in fact a pre-requisite since any reaction that actually occurs must be both kinetically and thermodynamically favoured. It is obvious that of all the reactions allowed by thermodynamics, the quickest will be the one first observed. The corollary is that unless the reaction which is allowed by thermodynamics but least favoured happens to be the quickest, it can never take place. Ostwald's rule has the logical consequence that in any sequence of precipitation the thermodynamically least-stable precipitates either appear first or not at all.

In practice, thermodynamic modelling of natural aqueous solutions generally reveals that many of them are highly supersaturated. However, such solids do not form (precipitate) because of kinetic constraints. Ostwald's rule can thus be applied very usefully to predict which solid(s) are most likely to precipitate by considering only those that lie fairly near to saturation, with a reasonable probability that the solid first to appear will be amongst the least supersaturated.

The solids listed in Table 4 are therefore those liable to precipitate in blood plasma according to Ostwald's rule. Most of them seem to be reasonable predictions, in line with general biochemical expectations. On the other hand, they are evidently limited in a quantitative sense by the twin issues that crop up throughout this paper - scarcity of the necessary thermodynamic parameters and difficulty in describing the effects of the electrolyte medium. For instance, the predicted solubility in blood plasma of sodium bicarbonate (given in Table 4) seems too low; at 25 °C and the relevant ionic strength in NaCl(aq) solution, the measured solubility is about 1 $mol.kg^{-1}$ (Linke, 1965). Difficulty in characterising the sodium bicarbonate ion pair is undoubtedly a contributing factor here but it is noteworthy that the JPD database has solubility products for sodium bicarbonate only at infinite dilution, i.e. not in NaCl(aq) or even in another electrolyte medium. It is indeed remarkable how few conditional equilibrium constants for reactions of solids appear to be available in well-defined

electrolyte solutions at finite concentrations and 25 °C. This is true in particular also for calcite and aragonite, which is relevant to the discussion below.

Weak chemical interactions

Of paramount concern in the seawater model are the effects of weak complexation. At the heart of this problem lies uncertainty regarding the fundamental nature of the specific interactions occurring in strong electrolytes (Marcus and Hefter, 2006). Although difficult, mostly impossible, to characterise unequivocally, these interactions become critical in modelling highly concentrated solutions. As a result, the accuracy of generalised chemical modelling of multicomponent solutions is compromised much more than is sometimes appreciated.

Under Scenarios A - D, the chloride complexes of sodium, magnesium, calcium, potassium and strontium are all calculated to be minor but significant species (e.g. $NaCl^0$ and $MgCl^+$ ion pairs predicted to account for 6% and 1% respectively of total chloride). While these calculated percentages are fairly small they depend heavily on the respective binding constants, **none of which can be properly quantified in any general sense**. Unfortunately, the physicochemical interactions they represent become significant at higher electrolyte concentrations. The situation is even worse with $CaCl^+$ which is found to account for 22% of the total calcium concentration, despite merely having lgK = –0.12 at I = 0.65 M. It is therefore necessary to treat these calculated speciation distributions with great caution and to draw only very high-level conclusions from them. The exact percentages evidently have little or no significance.

So difficult is experimental characterisation of the chloride species of the alkali and alkali earth metal ions (Marcus and Hefter, 2006) it is reasonable to think that they do not exist or that they should, at least, be ignored (May, 2006). The results obtained for Scenarios A' - D' explore this proposition. Although the saturation indices given in Table 12 then 'improve' significantly (meaning they move closer to the values reported in the literature for seawater) there is much less 'success' with the other carbonate-related quantities in Table 12. Evidently this is because these reactions are not so closely linked to chloride concentrations. They are, on the other hand, linked intimately with sodium ion binding to carbonate and bicarbonate, where the interactions are much stronger than those with chloride and thus not as lightly ignored.

The problem is that it is impossible to distinguish objectively (Marcus and Hefter, 2006) between what is sometimes called 'ion pairing' (to indicate weak affinity) and 'complexation' (to indicate something stronger). Whether such effects can be realistically represented by an equilibrium constant in a multicomponent electrolyte solution is still an open question. Even though such binding 'constants' can be estimated in certain carefully-selected swamping electrolytes, it is not experimentally possible to show, even under these controlled circumstances, that they are actually unchanging.

Calculating some solubilities in seawater-like solutions may thus be improved quantitatively by ignoring selected chloride interactions as shown in Table 12. However, such expedience is known to lead to its own problems in other contexts. A not-too-mysterious example can be seen in Table 8 where omission of the (calcium) chloride ion pair (Scenarios C1', C2' and D') and the consequent percentage increase in free chloride causes a *decrease* in the predicted percentage of the Cu+1_Cl-1(2) species.

The bottom line is that physicochemical interactions between species influence the reactions in ways that are currently difficult or impossible to characterise in terms of competing equilibria. Such problems are especially evident in (a) pure / neat solutions of single electrolytes and (b) electrolyte solution mixtures at moderate or high concentrations when chemical reactions taking place in the solution are not controlled by a predominant and inert background electrolyte. Nowhere is this dire issue more difficult than in the prediction of solubilities.

As long as one deals with just **one** swamping background electrolyte at fixed concentration, the activity coefficients of all minor species in solution remain constant regardless of weak interactions and so do the experimentally-determined equilibrium quotients of reactions (the so-called 'conditional equilibrium constants'). This is in accord with the practice popularised by Biedermann and Sillén (1953) of measuring equilibria in a supporting electrolyte medium. Moreover, these conditional constants will change in well-behaved ways as functions of non-specific quantities such as the formal ionic strength (May, 2000). On the other hand, since the type and extent of these indeterminate weak ionic interactions alters from one background electrolyte to another, the relationship between conditional equilibrium constants measured in different salt solutions of the same ionic strength is quantitatively unpredictable. Under these circumstances, most evident in highly concentrated solutions, the concept of ionic strength as a master variable determining activity coefficients starts to fail. Indeed, with increasing reactivity of the background electrolyte and/or increasing concentration, these confounding effects can become quantitatively ruinous.

Unless the species in a reaction are entirely and unrealistically indifferent to the specific nature of the background electrolyte ions - meaning that the same measured value is obtained for a wide variety of supposedly inert salts - errors arising from any mismatch in the medium progressively manifest themselves in all speciation models as the concentrations rise and/or the reactivity of the background electrolytes increase. The SIT equation is recommended by IUPAC (Powell et al., 2009) to deal with these specific ion interactions - which is an approach similar to that followed by JESS (May, 2000) – but no one has properly investigated how such equations perform in multicomponent electrolyte solutions. Similarly, attempts (Helgeson et al., 1981, p. 1284) to use the 'true ionic strength' (meaning the ionic strength calculated at equilibrium when, for instance, neutral species are formed by ion pairing) are not useful, and may even be deceptive, because the conditional equilibrium constants used in speciation calculations are quantified in terms of 'formal' / 'stoichiometric' ionic strength under the assumption of complete ionic dissociation.

To address the issues described above, major efforts have been made to couple chemical speciation calculations with 'ion-interaction' models such as those based on the Pitzer equations. These so-called 'hybrid' approaches aim to retain the usefulness of speciation modelling derived from the vast number of measured equilibrium constants while simultaneously taking advantage of the ability of

ion-interaction frameworks to characterise the thermodynamics of bulk electrolyte solutions. Again, unfortunately, rather little has been done to test the effectiveness of such combined models, particularly in the context of solubility predictions. Neither has there been any significant examination of the theoretical incompatibilities inherent in these hybrid calculations (Grenthe et al., 1997, p. 356), nor attention paid to the obviously artificial outcome of the convention employed to split the cation and anion activity coefficients.

Beyond even these concerns, the most critical issue for any hybrid model is that it will ultimately suffer the same constraints as all ion-interaction frameworks (Rowland et al., 2014) due to their empirical nature and a combinatorial explosion of correlated parameters. Finding ways to overcome these limitations is therefore a critical challenge ahead.

Data uncertainty

Modelling reliability depends first and foremost on the availability of quality data - experience has shown that where the distributions came from experimentally-well-characterised species the model's results proved resilient over time but where there were only a few underlying literature values little confidence in longevity was warranted. There are especially good reasons to be suspicious of predominant species that are predicted by models from just one or two reported equilibrium constants - by their very nature, such chemical systems, e.g. aluminium(III)-phosphate (Alliey et al., 1995), tend to be prone to artefact and extrapolation errors. Another example is provided by the demise of ' Zn+2_Citric-3_Cys-2' in Table 3 where a single literature datum on which the original calculation was based (Ramamoorthy and Manning, 1974) simply became discredited (Berthon et al. 1978). This is in sharp contrast to frequently investigated systems where 'the democracy of observers' generally works to weed out major systematic errors.

Nevertheless, it is easy to underestimate how bad the bedrock of thermodynamic parameters can be, even today. Careful comparisons between measured thermodynamic quantities tend to reveal truly remarkable differences when experimentalists are strictly independent (Rowland and May, 2013). For a further example, Figure 7 in Millero (2007) clearly shows that the values of some equilibrium constants critical to seawater modelling are hardly known to better than the first decimal place (log units). Careless errors are also common throughout the chemical literature (see e.g. Lewis and Wallace, 2014). Avoiding such artefacts is therefore an important but not always easy task. How to overcome the profound and pervasive corruption of thermodynamic data that can be caused by deeply-embedded errors - see e.g. Filella and May (2003) for some examples - or sharp changes in preferred values - see e.g. Akilan et al., (2014) for one such recent change - or revised fundamental constants is a largely unsolved problem, affecting JESS and all other large assemblages of thermodynamic data.

Thermodynamic modelling under these treacherous conditions is difficult; experience with JESS has shown that it needs more than subjective opinion, no matter how authoritative. On the other hand, automatic mechanisms are no substitute for the comprehensive and critical assessments of data by teams of experts such as are assembled under the auspices of IUPAC, e.g. Powell et al., (2013) and

the OECD Nuclear Energy Agency. More automation of the kind that JESS has pioneered is nonetheless essential because critically-evaluated datasets are also severely limited, both in scope and in keeping up-to-date, by the enormous workload involved.

Strengths of JESS

The main objective advantages of the JESS speciation software are calculation versatility, various convenient means for the output of modelling results, database accessibility - visible and modifiable literature sources and database verification mechanisms, including automatic thermodynamic consistency checking.

Regarding calculation versatility, the advantages over the facilities of some other packages fall into two categories: (a) explicit control of the ways the modelling equations can be set up and solved and (b) conceptually simpler ways of achieving particular chemical modelling outcomes.

Examples of (a) include:

- Partitioning of chemical reactions to model metastable chemical systems in which certain reactions are kinetically-constrained, e.g. by decoupling the chemical reactions of sulfate from other sulfur-containing species such as sulfide (which can co-exist without coming to equilibrium for periods much greater than the timescales of interest)
- Facile modification of the mathematical basis functions used to set up and solve the modelling equations, thus ensuring that difficulties in achieving convergence can *always* be overcome (provided only that the modelling Scenario is realistic)
- Direct imposition of pH and Eh values - so that they do not always need to be determined by an iterative solution of the modelling equations ('reconciliation' in one terminology); while the two approaches should yield identical answers, the need to evaluate such quantities can be problematic. For example, in Scenario A of the seawater model all speciation-focussed calculations would struggle to deal properly with the ratios of iron(II) and iron(III) species at such very low concentrations, a difficulty that does not arise if the Eh can be specified rather than calculated. Similar numerical difficulties also crop up in such calculations as a result of small errors in the relevant equilibrium constants, difficulties that are not manifest to the same extent if the Eh can be fixed as an independent variable. This is a good illustration of the principle that thermodynamic consistency should always be judged relative to experimental uncertainty and not as an absolute criterion.

Examples of (b) include:

- Facilities to impose on the chemical speciation calculation independently-ascertained equilibrium constants determining the formation of non-basis species (such as complexes and ion pairs)

- User discretion to impose, or not to impose, the charge balance as an explicit mathematical constraint on the speciation calculations.

A special feature of JESS is its ability to produce a complete, stand-alone FORTRAN program for any particular chemical equilibrium model that has been developed. These automatically-generated, compact FORTRAN codes (called 'Mini-Models') are application specific, with all the chemical relationships and thermodynamic parameters fixed in advance by setting up an ordinary JESS model calculation. In this way the overheads normally incurred by JESS for the sake of generality are avoided. So Mini-Models perform speciation calculations with exceptional efficiency and speed, which can be exploited by linking them directly into process simulators.

Weaknesses of JESS

One frequently perceived disadvantage of the JESS package is that the user-interface is command-driven and operates solely from within one or more DOS-windows. Without a 'point and click' mechanism, new users must acquire significant expertise in the conventions used by the package to manipulate data and to drive the modelling procedure. Some dedicated effort is generally required for a user to become familiar with the necessary background information and protocols. Details can be found in the JESS Primer, which can be downloaded from the website (http://jess.murdoch.edu.au).

A significant technical omission of JESS is the lack of pressure dependence in calculated speciation distributions (even though the database does record some thermodynamic parameters as functions of pressure). Insufficient experimental data - an unfortunately common occurrence - can also prove troublesome in estimating the effects of temperature change on solutions at moderate or high ionic strength. While JESS includes a number of simple models for chemical adsorption, this capability is now deprecated. Likewise, there is no capability for dealing with ion exchange or solid solutions of varying composition.

Future needs

As is evident from the modelling exercises described in this paper, although checks for automatic consistency help, harmonising large quantities of thermodynamic data for chemical reactions remains a formidable and onerous task. Overcoming the problems due to data uncertainty is still the most intractable problem hindering general thermodynamic modelling of aqueous solutions. In essence, the existing JESS facilities are still not powerful enough to exploit all the information content available. Rather than treating reactions individually, global mechanisms must be encoded, not only to produce better models but also to make more sustainable the processes of data assessment, of redundancy elimination and of averaging.

Only through carefully-designed optimisation strategies and better automatic data processing can we cope better with continuous changes to our thermodynamic databases, be these from new measurements, corrections or critical re-assessments. While inconsistencies in the JPD database

have been reduced considerably since it first appeared (May and Murray, 1991b), obtaining reliable thermodynamic parameters for many reactions in aqueous solution remains problematic. It is worth noting that a tolerance of 0.5 log units for differences arising from linear combinations of reactions is still being taken by JESS as 'satisfactory' (see Appendix). Furthermore, in the present work, it was only possible to obtain reasonable consistency amongst all reaction subsets for each relevant element on their own (in water) and in certain simple combinations; this was not possible for the chemistry associated with more complicated combinations (such as arises with aluminosilicates), where the default modelling process can only ensure that those best-characterised reactions are used in preference to others. This is because complicated substances, such as aluminosilicates, often involve numerous reactions in linear combination.

Somewhat predictably therefore, it seems that significant progress will require greater theoretical understanding and even more experimental data than have been accumulated to date. However, given the hard task these days in getting the resources needed for experimental work in solution chemistry, the burden must likely fall most heavily on theoretical advances and better ways of processing the data than are already available. JESS facilities are currently being developed to express all the aqueous solution property data which have been compiled for chemical reactions and physicochemical quantities (including solubilities) combined together on a uniform and readily processable basis. There are currently about 700,000 values so assembled in JESS and it is anticipated that this number will double in the foreseeable future. We envisage that we will then be able to refine model parameters more effectively and more comprehensively than has so far been the case.

CONCLUSION

It is fair to say that, in the conception of JESS, an early hope of producing a computing facility that would allow inexpert users to construct error-free chemical speciation models easily, rapidly and reproducibly has not been achieved. Perhaps that goal was just too optimistic. It turns out instead that JESS helps solution chemists with a reasonable degree of expertise to develop robust and reliable chemical speciation models through a straightforward, but usually iterative, process. Importantly, the process is designed to expose the assumptions and uncertainties that inevitably underpin these kinds of calculation.

At present, all models predicated on a chemical speciation paradigm (i.e. using a so-called 'ion-association' framework), exhibit a common set of fundamental strengths and weaknesses: they are good at describing strong chemical effects between components in solution but poor in dealing with weaker interactions. Consequently, large differences can occur between such models regarding particular species concentrations and percentage distributions.

Fortunately, the adverse consequences of these differences in percentage distributions can often be avoided by adopting appropriate strategies. Ion-association packages are then well able to simulate, in the most general and comprehensive way, the main effects of chemical change in key solution variables such as pH and redox potential. Thus, the chemical speciation framework proves valuable in practical applications such as modelling kidney stone formation in urine (Pak et al., 2009) and in

hydrometallurgical leaching (Cifuentes et al., 2006) where it is unequalled in helping to understand the complicated chemistry involved. No other theoretical framework in solution chemistry has such reach.

For practical purposes it is modelling skill, chemical knowledge and familiarity with the modelling software that are the keys, rather than the precise capabilities of the computer programs and databases employed. As the co-founder of JESS, Kevin Murray, taught us: "The art of scientific modelling is ... finding the right questions to the answers; not being afraid to make assumptions; how you do it, not what you do." (May et al., 2014). A regrettably frequent mistake of would-be modellers is to expect chemical speciation calculations to deliver them a single (exact) solution. For natural aqueous systems in general (as opposed to laboratory solutions) this is beyond current capability and probably always will be. To the contrary, the prime objective of scientific modelling should be to gain insight. For this reason JESS has been developed as a research tool [http://jess.murdoch.edu.au] not just a speciation program code. And, in this sense, the JESS project that commenced thirty years ago seems to have achieved its founding ambitions. However, considerable work must still be done to produce accurate models of complex aqueous solutions over the whole range of desired conditions for composition, concentration, temperature and pressure. This includes a greater understanding of the fundamental interactions in solution giving rise to observed thermodynamic properties, as well as an even greater effort directed at stabilising the underlying thermodynamic databases. Only in this way can aqueous solution modelling become less susceptible to measurement error and to the so-often erratic nature of experimental results published in the literature.

ACKNOWLEDGEMENTS

Thanks go to my employer, Murdoch University, for the salary and other resources provided in part to support JESS since 1985. Thanks also go to Eric May (University of Western Australia) and my Murdoch colleagues Glenn Hefter, Erich Königsberger and Darren Rowland for their constructive comments on this manuscript and for many interesting conversations about thermodynamics and solution chemistry. It is a pleasure for me to acknowledge that JESS could not have been conceived or implemented without the inspiration and enormous effort of Kevin Murray (formerly of The CSIR, South Africa).

APPENDIX

The following selected extracts from a simple calculation performed through the GEM stages illustrate the general approach to speciation modelling taken by JESS and the kinds of output available. Explanatory comments are italicised.

Program TELSUB

Allows users to control the dataset being extracted from the main JESS database (JPD). This defines, through the following lists, the compositional domain of the speciation calculation(s) to be performed.

```
1.  Elements
2.  Primitives
3.  Advanced Facilities
```

'Primitives' are chemical species with simple I/O (input/output) symbols, e.g. H+1, Cu+2, EDTA-4. 'Composites', in contrast, are species with symbols comprising two or more primitives separated by an underscore character, e.g. H+1_CO3-2, Cu+2_EDTA-4, Fe+3_EDTA-4_OH-1. I/O symbols are for the user interface only; no distinction occurs elsewhere – all species are characterised internally by their molecular formula and electronic charge, together with other information where available such as their CAS Registry number.

Program VEWLGK

Allows users to inspect the results of the process that has led to thermodynamic consistency.

```
 1.  List of Selected Primitive Species
 2.  All Identified Species
 3.  Reactions (with info content, reliability, lgK)
 4.  Default MBU Basis Species Set Order
 5.  Default Variable Basis Species Set Order
 6.  Thermodynamic Basis Species Set
 7.  Reactions forming LGK non-basis species with a given primitive
 8.  Linear combinations of reactions from the LGK-stage
 9.  Rogues' gallery
10.  Best lgK/dG and UCCs for any possible reaction
11.  Conditions, Methods and Assumptions
12.  Advanced Facilities

Select one (1-12) <9> :

>>  30 reactions compared for:  Pre = 1 bar  Temp = 25 C  IStr = 0.000 M

O2 + 4<e-1> + 2<H2O> = 4<OH-1>  (Rxn 96) IC=3
------------------------------------------------------------------------------
>>  lgK diff. = 2.96    Rxn. 96 lgK = 27.13    Lin. comb. lgK = 30.09
------------------------------------------------------------------------------
(-2) H+1 + OH-1 = H2O  (Rxn 1) IC=881
(-) O2(g) = O2  (Rxn 31) IC=12
(+2) 0.5<H2(g)> + 0.5<O2(g)> + e-1 = OH-1  (Rxn 25) IC=85
(+2) H+1 + e-1 = 0.5<H2(g)>  (Rxn 23) IC=999

>>  1 reaction with a difference in lgK > 0.5
```

IC (Information Content) is based mainly on the number of equilibrium constants available for the given reaction under particular conditions but it includes other factors such as the reliability (weight) that is individually assigned to equilibrium constants in the reaction database. ICs vary from 0 to 999.

The full set of equilibrium constants (lgK) and reaction enthalpies (dH) for the reactions shown here can be seen on the JESS website [http://jess.murdoch.edu.au].

Program VEWBAS

Displays results of the process which determines the variable and MBU basis sets. The concentrations of the variable basis species are the algebraic unknowns found by solving the mass balance equations. The MBUs (mass balance units) are the species for which total concentrations are (generally) known. These two basis sets can be the same but they generally differ, a feature that ensures the mass balance equations can always be formulated in a way that is numerically well-conditioned. The UCCs (Unconditional Correction Coefficients) are the parameters of the functions used to calculate the conditional equilibrium constants for given temperature and ionic strength – see May (2000) for further detail.

```
     1.  Scan through UCCs
     2.  UCCs for specified species
     3.  lgK for specified species at given conditions

Select one (1-3) <1> :

H+1_CO3-2 + H+1 = H+1(2)_CO3-2  (Rxn 8) IC=999

UCCs: lgK0=6.349; dH0=-6339; dCp0=313.7; GugC=-0.1933; GugM=0.000742

P scores (bar):  1,888(<100)  0(<2000)  0(<5000)  0(>5000)
T scores (C): 359(<20) 590(<34) 556(<95) 335(<200) 225(<300) 0(>300)
I scores (M): 865(<.01) 207(<.20) 317(<.85) 178(<1.8) 169(<3.1) 152(>3.1)

……

Ca+2 + H+1_CO3-2 = H+1 + Ca+2_CO3-2

(+) Ca+2 + CO3-2 = Ca+2_CO3-2  (Rxn 6) IC=836
(-) H+1 + CO3-2 = H+1_CO3-2  (Rxn 4) IC=933

UCCs: lgK0=-7.156;  dH0=3.007E+4;  dCp0=0.1125;  GugC=-0.7375;  GugM=-0.0001447
```

'Scores', like IC (Information Content), are counters based on the number of equilibrium constants available for the given reactions under particular conditions of pressure (P), temperature (T) and Ionic Strength (I). These counters are used in the ordering of the reactions and species, which in turn determines (by Gaussian elimination) the default basis sets and the selected sets of linear combination of reactions which are to be used in the subsequent speciation calculations.

Program VEWQED

Allows users to inspect interactively the results of the speciation calculation.

```
You may view the following at the current scan point
```

```
     1.  Speciation of Specified Elements
     2.  Speciation of Specified Primitives
     3.  Total Proton or Electron Concentrations
     4.  Individual Species Concentrations
     5.  Phase Distribution of Specified Elements
     6.  Saturation State of Solids
     7.  Partial Pressure of Gases
     8.  Charge Balance/Ionic Strength
     9.  Water Quality Indexes
    10.  Solution to Equations
    11.  Data Quality

Select one (1-11) <1> :

The predominant species containing the elements you specify will be displayed

Number of predominant species required <5> :

Elements : C

   C       mg/L C or
moles/L    microg/L *    %    Species
-------    ----------   --    -------
0.0008062     9,684 *   81    H+1_CO3-2
0.0001803     2,166 *   18    H+1(2)_CO3-2
1.251E-5        150 *    1    Ca+2_H+1_CO3-2
5.550E-7          7 *    0    Ca+2_CO3-2
3.618E-7          4 *    0    CO3-2
--------   ----------   --    -------
0.001000         12    100    Total C in above species
```

Program TELOUT

Allows users to specify various possible kinds of output from the QED speciation calculation (called 'determinands'). The required values can subsequently be put into a text file (comma delimited ASCII) suitable for importation into spreadsheets. The output from large, multi-dimensional scans can thus be streamed conveniently to produce graphical representations of the calculated results, such as plots of pH-Eh and other speciation profiles.

```
No.   Determinand                                          Info required

 1    Temperature [C]
 2    Ionic Strength [M]
 3    -lg[H+1]
 4    Oxidation Potential [mV]
 5    lg(Equilibrium Constant) ........................... Species Symbol
 6    Concentration [M] of species ....................... Species Symbol
 7    lg(SI) of solid .................................... Solid Species Symbol
 8    lg(Partial Pressure of Gas) ........................ Gas Species Symbol
 9    Total Concentration [M] of element ................. Element Symbol
10    Single Predominant Species of element .............. Element Symbol
11    Single Predominant Species of primitive ............ Primitive Symbol
13    Total Aqueous Concentration [M] of element ......... Element Symbol
14    Total Solid Concentration [M] of element ........... Element Symbol
16    Total Liquid Concentration [M] of element .......... Element Symbol
```

```
17    Total Aqueous Concentration [M] of primitive ....... Primitive Symbol
18    Total Solid Concentration [M] of primitive ......... Primitive Symbol
19    Total Liquid Concentration [M] of primitive ........ Primitive Symbol
20    Total Proton Concentration [M]
21    Total Electron Concentration [M]
22    Data Quality ..................................... Element Symbol/any
23    Charge Imbalance [meq/L]
24    % of Negative or Positive Charge not Balanced
25    No. of Supersaturated Solids ..................... Element Symbol/any
26    Single Predominant Aqueous Species of element ...... Element Symbol
27    Single Predominant Aqueous Species of primitive .... Primitive Symbol
28    [C(i)xdH(i)] non-Variable Basis Species ............ Species Symbol
29    SUM[C(i)xdH(i)] all non-Variable Basis Species
30    % Charge Imbalance (Std. Methods)
31    ln(Concentration) [M] of species ................... Species Symbol
32    Total Adsorbed Concentration [M] of element ........ Element Symbol
33    Number of Newton-Raphson Iterations
34    Total negative charge [meq/L anions]
35    Total positive charge [meq/L cations]
36    Pressure [bar]
37    Internal JESS Completion Status Flag
38    H2O density [kg/m3]
39    Debye-Hueckel A constant
```

Program PRTOUT

Generates a comma-delimited ASCII file of results. For example, after QED a scan specifying 6 speciation calculations are to be performed between pH 7 and pH 8, the following specification of determinands in TELOUT:

```
3
6 Ca+2_CO3-2
7 Ca+2_CO3-2_(Calc.,s)
```

produces an output (CSV) file that looks like this:

```
"pH","CaCO3","lgSI(Calcite)"
7.0000,5.5500E-7,-1.0098
7.2000,9.4061E-7,-0.78068
7.4000,1.5576E-6,-0.56164
7.6000,2.5366E-6,-0.34983
7.8000,4.0818E-6,-0.14323
8.0000,6.5077E-6,0.059344
```

Of course, 2-D predominance diagrams (e.g. for pH-Eh) typically involve many thousands of scan points, not just six as in this example. The OUT programs are designed to facilitate collection and analysis of such large arrays of chemical information.

REFERENCES

Akilan, C., May, P.M. and Hefter, G. 2014. A potentiometric study of the association of copper(II) and sulfate ions in aqueous solution at 25 °C. J. Solution Chem., **43,** 885-892.

Alliey, N., Biron, F., Venturini-Soriano, M., Berthon, G., 1995. A Reinvestigation of Aluminium-Phosphate Equilibria under Physiological Conditions. Preliminary Results.. *J. Inorg. Biochem*. **59**, 241-241.

Anderson, G.M., Castet, S., Schott, J., Mesmer, R.E., 1991. The Density Model for Estimation of Thermodynamic Parameters of Reactions at High Temperatures and Pressures. *Geochim. Cosmochim. Acta* **55**, 1769-1779.

Berthon, G., 1995, Computer-based Studies on Trace Metal Excretion and Needs in Total Parenteral Nutrition. in: Berthon, G. (Ed.) Handbook of Metal-Ligand Interactions in Biological Fluids. Bioinorganic Chemistry. Volume 2, Marcel Dekker Inc., N.Y., U.S.A., pp. 1240-1251.

Berthon, G., Hacht, B., Blais, M.-J., May, P.M., 1986. Copper-Histidine Ternary Complex Equilibria with Glutamine, Asparagine and Serine. The Implications for Computer-simulated Distributions of Copper(II) in Blood Plasma. *Inorg. Chim. Acta* **125**, 219-227.

Berthon, G., Lambs, L., 1995, Speciation Studies in Relation to the Bioavailability and Drug Activity of Tetracyclines. in: Berthon, G. (Ed.) Handbook of Metal-Ligand Interactions in Biological Fluids. Bioinorganic Chemistry. Volume 2, Marcel Dekker Inc., N.Y., U.S.A., pp. 1252-1265.

Berthon, G., May, P.M., Williams, D.R., 1978. Computer Simulation of Metal-Ion Equilibria in Biofluids. Part 2. Formation Constants for Zinc(II)-Citrate-Cysteinate Binary and Ternary Complexes and Improved Models of Low-Molecular-Weight Zinc Species in Blood Plasma. *J. Chem. Soc., Dalton Trans.* 1433-1438.

Biedermann, G., Sillén, L.G., 1953. Studies on the Hydrolysis of Metal Ions. IV. Liquid Junction Potentials and Constancy of Activity Factors in $NaClO_4$-$HClO_4$ Ionic Medium. *Arkiv Kemi* **5**, 425-440.

Blais, M.-J., Berthon, G., 1982. Zinc-Citrate Interactions in Blood Plasma. Quantitative Study of the Metal Ion Equilibria in the Zinc-Citrate-Histidinate, -Glutaminate, and -Threoninate Systems and Computer Simulation of the Ability of Citrate to Mobilize the Low Molecular Weight Fraction of Zinc. *Inorg. Chim. Acta* **67**, 109-115.

Brumas, V., Alliey, N., Berthon, G., 1993. A New Investigation of Copper(II)-Serine, Copper(II)-Histidine- Serine, Copper(II)-Asparagine, and Copper(II)-Histidine-Asparagine Equilibria under Physiological Conditions, and Implications for Simulation Models Relative to Blood Plasma. *J. Inorg. Biochem.* **52**, 287-296.

Butler, J.N., 1992. Alkalinity Titration in Seawater: How Accurately Can the Data be Fitted by an Equilibrium Model. *Mar. Chem.* **38**, 251-282.

Byrne, R.H., Kump, L.R., Cantrell, K.J., 1988. The Influence of Temperature and pH on Trace Metal Speciation in Seawater. *Mar. Chem.* **25**, 163-181.

Cifuentes, L., Casas, J.M., Simpson, J., 2006. Temperature Dependence of the Speciation of Copper and Iron in Acidic Electrolytes. *Chem. Eng. Res. Design* **84**, 965-969.

Clegg, S.L., Whitfield, M., 1995. A Chemical Model of Seawater Including Dissolved Ammonia and the Stoichiometric Dissociation Constant of Ammonia in Estuarine Water and Seawater from -2 to 40 °C. *Geochim. Cosmochim. Acta* **59**, 2403-2421.

Cole, A., Furnival, C., Huang, Z.-X., Jones, D.C., May, P.M., Smith, G.L., Whittaker, J., Williams, D.R., 1985. Computer Simulation Models for the Low-molecular-weight Complex Distribution of Cadmium(II) and Nickel(II) in Human Blood Plasma. *Inorg. Chim. Acta* **108**, 165-171.

Crea, F., Giacalone, A., Gianguzza, A., Piazzese, D., Sammartano, S., 2006. Modelling of Natural and Synthetic Polyelectrolyte Interactions in Natural Waters by using SIT, Pitzer and Ion Pairing Approaches. *Mar. Chem.* **99**, 93-105.

Darn, S.M., Sodi, R., Ranganath, L.R., Roberts, N.B., Duffield, J.R., 2006. Experimental and Computer Modelling Speciation Studies of the Effect of pH and Phosphate on the Precipitation of Calcium and Magnesium Salts in Urine. *Clin. Chem. Lab. Med.* **44**, 185-191.

Darwish H.M., Schmitt R.C., Cheney J.C. and Ettinger M.J., 1984. Mobilization of Copper(II) from Plasma Components and Mechanism of Hepatic Copper Transport. *Am. J. Physiol.*, **246**, G72-G79.

Davidge, J., Williams, D.R., 2003. Chemical Speciation to Optimise the Efficiency of Mixed Blends for Complexing in Industry and Healthcare. Inorg. Chim. Acta **356**, 109-113.

Dayde, S., Brumas, V., Champmartin, D., Rubini, P., Berthon, G., 2003. Aluminum Speciation Studies in Biological Fluids. Part 9. A Quantitative Investigation of Aluminum(III)-Glutamate Complex Equilibria and their Potential Implications for Aluminum Metabolism and Toxicity. *J. Inorg. Biochem.* **97**, 104-117.

Dickson, A.G., Millero, F.J., 1987. A Comparison of the Equilibrium Constants for the Dissociation of Carbonic Acid in Seawater Media. *Deep-Sea Res.* **34**, 1733-1743.

Dickson, A.G., Whitfield, M., 1981. An Ion-association Model for Estimating Acidity Constants (at 25 °C and 1 atm Total Pressure) in Electrolyte Mixtures Related to Seawater (Ionic Strength < 1 mol kg^{-1} H_2O). *Mar. Chem.* **10**, 315-333.

Dickson, A.G., Whitfield, M., Turner, D.R., 1981. Concentration Products: Their Definition, Use and Validity as Stability Constants. *Mar. Chem*. **10**, 559-565.

Filella, M., 1995, Speciation Models in Relation to Pharmaceutical Activity. in: Berthon, G. (Ed.) Handbook of Metal-Ligand Interactions in Biological Fluids. Bioinorganic Chemistry. Volume 2, Marcel Dekker Inc., N.Y., U.S.A., pp. 1240-1251.

Filella, M., May, P.M., 2003. Computer Simulation of the Low-Molecular-Weight Inorganic Species Distribution of Antimony(III) and Antimony(V) in Natural Waters. *Geochim. Cosmochim. Acta* **67**, 4013-4031.

Garrels, R.M., Thompson, M.E., 1962. Chemical Model for Seawater at 25 °C and 1 Atmosphere Total Pressure. *Am. J. Sci.* **260**, 57-66.

Gautier, Q., Bénézeth, P., Mavromatis, V. and Schott, J. 2014. Hydromagnesite Solubility Product and Growth Kinetics in Aqueous Solution frrom 25 to 75 °C. *Geochim. Cosmochim. Acta*, **138**, 1-20.

Grenthe, I., Plyasunov, A.V., Spahiu, K., 1997. Estimation of Medium Effects on Thermodynamic Data. in: Grenthe, I., Puigdomenech, I. (Eds.), Modelling in Aquatic Chemistry., Nuclear Energy Agency, OECD, Paris, France, pp. 325-426.

Hallman, P.S., Perrin, D.D., Watt, A.E., 1971. The Computed Distribution of Copper(II) and Zinc(II) Ions Among Seventeen Amino Acids Present in Human Blood Plasma. *Biochem. J.* **121**, 549-555.

Harvie, C.E., Weare, J.H., 1980. The Prediction of Mineral Solubilities in Natural Waters: The Na-K-Mg-Ca-Cl-SO_4-H_2O System from Zero to High Concentration at 25 °C. *Geochim. Cosmochim. Acta* **44**, 981-997.

Helgeson, H.C., Kirkham, D.H., Flowers, G.C., 1981. Theoretical Prediction of the Thermodynamic Behavior of Aqueous Electrolytes at High Pressures and Temperatures. IV. Calculation of Activity Coefficients, Osmotic Coefficients and Apparent Molal and Standard and Relative Partial Molal Properties to 600 °C and 5 Kb. *Am. J. Sci.* **281**, 1249-1516.

Jackson, G.E., 1995a, Speciation Approach to the Design of Antiinflammatory Copper Complexes. in: Berthon, G. (Ed.) Handbook of Metal-Ligand Interactions in Biological Fluids. Bioinorganic Chemistry. Volume 2, Marcel Dekker Inc., N.Y., U.S.A., pp. 1228-1239.

Jackson, G.E., 1995b, Computer-aided Approach to Imaging Contrast Agents. in: Berthon, G. (Ed.) Handbook of Metal-Ligand Interactions in Biological Fluids. Bioinorganic Chemistry. Volume 2, Marcel Dekker Inc., N.Y., U.S.A., pp. 1281-1290.

Königsberger, E., Königsberger, L.-C., Gamsjager, H., 1999. Low-temperature Thermodynamic Model for the System Na_2CO_3-$MgCO_3$-$CaCO_3$-H_2O. *Geochim. Cosmochim. Acta* **63**, 3105-3119.

Lewis, E. and Wallace, D., 2014. http://cdiac.esd.ornl.gov/oceans/co2rprtnbk.html

Lide, D.R., 2005. CRC Handbook of Chemistry and Physics, 86th Edn., A Ready-Reference Book of Chemical and Physical Data, Taylor & Francis, Boca Raton, USA.

Linke, W.F., 1965. Solubilities, Inorganic and Metal-Organic Compounds, Vol. 2, A Revision and Continuation of the Compilation Originated by Atherton Seidell, Ph.D., U.S. National Institutes of Health, American Chemical Society.

May, P.M., 1990. Agents Acting on Ion Transport - Passive Systems. Comprehensive Medicinal Chemistry, Vol 2., pp. 175-192.

May, P.M., 2000. A Simple, General and Robust Function for Equilibria in Aqueous Electrolyte Solutions to High Ionic Strength and Temperature. *J. Chem. Soc., Chem. Commun.* 1265-1266.

May, P.M., 2006. Improved Thermodynamic Calculations for Concentrated Mixed Electrolyte Systems Including Ion Pairing (or the Absence of It). *Marine. Chem.* **99,** 62-69.

May, P.M., Bulman, R.A., 1983. The Present Status of Chelating Agents in Medicine. *Prog. Med. Chem.* **20**, 225-336.

May, P.M., Linder, P.W., Williams, D.R., 1977. Computer Simulation of Metal-ion Equilibria in Biofluids: Models for the Low-Molecular-Weight Complex Distribution of Calcium(II), Magnesium(II), Manganese(II), Iron(III), Copper(II), Zinc(II), and Lead(II) Ions in Human Blood Plasma. *J. Chem. Soc., Dalton Trans.* 588-595.

May, P.M., Murray, K., 1991a. JESS, A Joint Expert Speciation System - I. Raison d'Etre. *Talanta* **38**, 1409-1417.

May, P.M., Murray, K., 1991b. JESS, A Joint Expert Speciation System - II. The Thermodynamic Database. *Talanta* **38**, 1419-1426.

May, P.M., Murray, K., 2001. Database of Chemical Reactions Designed to Achieve Thermodynamic Consistency Automatically. *J. Chem. Eng. Data* **46**, 1035-1040.

May, P.M., Murray, K., Williams, D.R., 1985. The Use of Glass Electrodes for the Determination of Formation Constants. II. Simulation of Titration Data. *Talanta* **32**, 483-489.

May, P.M., Murray, K., Williams, D.R., 1988. The Use of Glass Electrodes for the Determination of Formation Constants - III. Optimization of Titration Data: The ESTA Library of Computer Programs. *Talanta* **35**, 825-830.

May, P.M., Rowland, D., Hefter, G., Königsberger, E., 2011. A Generic and Updatable Pitzer Characterization of Aqueous Binary Electrolyte Solutions at 1 bar and 25 °C. *J. Chem. Eng. Data* **56**, 5066-5077.

May, P.M., Rowland, D., Königsberger, E., Hefter, G., 2010. JESS, a Joint Expert Speciation System - IV: A Large Database of Aqueous Solution Physicochemical Properties with an Automatic Means of Achieving Thermodynamic Consistency. *Talanta* **81**, 142-148.

May, P.M., Rowland, D. and Murray, K., 2014. JESS Primer. Available at http://jess.murdoch.edu.au/docs/Jess_Primer_V82.pdf

Mehrbach, C., Culberson, C.H., Hawley, J.E., Pytkowicz, R.M., 1973. Measurement of the Apparent Dissociation Constants of Carbonic Acid in Seawater at Atmospheric Pressure. *Limnol. Oceanogr.* **18**, 897-907.

Millero, F.J., 2002. Sea Water as an Electrolyte. in: Gianguzza, A., Pelizzetti, E., Sammartano, S. (Eds.), Chem. Mar. Water Sediments., Springer-Verlag, Berlin, Germany, pp. 3-34.

Millero, F.J., 2007. The Marine Inorganic Carbon Cycle. *Chem. Rev.* **107**, 308-341.

Millero, F.J., Feistel, R., Wright, D.G.T., McDougall, T.J., 2008. The Composition of Standard Seawater and the Definition of the Reference-Composition Salinity Scale. *Deep-Sea Res., Part I* **55**, 50-72.

Orvig, C., Berthon, G., 1995, Speciation Studies in Relation to Aluminium Toxicity. in: Berthon, G. (Ed.) Handbook of Metal-Ligand Interactions in Biological Fluids. Bioinorganic Chemistry. Volume 2, Marcel Dekker Inc., N.Y., U.S.A., pp. 1266-1280.

Ostwald, W. 1897. Studien über die Bildung und Umwandlung fester Körper. 1. Abhandlung: Übersättigung und Überkaltung. *Z. Phys. Chem.* **22**, 289–330.

Pak, C.Y.C., Maalouf, N.M., Rodgers, K., Poindexter, J.R., 2009. Comparison of Semi-empirical and Computer Derived Methods for Estimating Urinary Saturation of Calcium Oxalate. *J. Urol.* **182**, 2951-2956.

Perrin, D.D., 1965. Multiple Equilibria in Assemblages of Metal Ions and Complexing Species: A Model for Biological Systems. *Nature (London)* **206**, 170-171.

Popov, K.I., Wanner, H., 2005. Stability Constants Data Sources: Critical Evaluation and Application for Environmental Speciation. *ACS Symp. Ser. 910*, 50-75.

Powell, K.J., Brown, P.L., Byrne, R.H., Gajda, T., Hefter, G., Leuz, A.-K., Sjoberg, S., Wanner, H., 2009. Chemical Speciation of Environmentally Significant Metals with Inorganic Ligands. Part 3. The Pb^{2+}-OH^-, Cl^-, CO_3^{2-}, SO_4^{2-}, and PO_4^{3-} Systems. *Pure Appl. Chem.* **81**, 2425-2476.

Powell, K.J., Brown, P.L., Byrne, R.H., Gajda, T., Hefter, G., Leuz, A.-K., Sjoberg, S., Wanner, H., 2013. Chemical Speciation of Environmentally Significant Heavy Metals with Inorganic Ligands. Part 5: The Zn^{2+} + OH^{-}, Cl^{-}, CO_3^{2-}, SO_4^{2-}, and PO_4^{3-} Systems (IUPAC Technical Report). *Pure Appl. Chem.* **85**, 2249-2311.

Ramamoorthy, S., Manning, P.G., 1974. Equilibrium Studies of Solutions Containing Pb^{2+} or Zn^{2+}, and Cysteine, Orthophosphate and a Carboxylic Acid. *J. Inorg. Nucl. Chem.* **36**, 1671-1674.

Read, D., Broyd, T.W., 1988. Report on Stage 1 of Project CHEMVAL/MIRAGE, Verification of Speciation Models, WS Atkins Eng. Services under CEC contract FI.IW0077.UK(H), Epsom, UK.

Rodgers, A., Allie-Hamdulay, S., Jackson, G.E., 2006. Therapeutic Action of Citrate in Urolithiasis Explained by Chemical Speciation: Increase in pH is the Determinant Factor. *Nephrol., Dial., Transplant.* **21**, 361-369.

Rodgers, A.L., Allie-Hamdulay, S., Jackson, G.E., 2007. JESS: What Can It Teach Us?. *AIP Conf. Proc.* **900**, 183-191.

Rolnick L.R., 1954. The Stability of Gypsum and Anhydrite in the Geologic Environment. Ph.D. Thesis, Massachusetts Institute of Technology.

Rowland, D., May, P.M., 2010. JESS, a Joint Expert Speciation System - V: Approaching Thermodynamic Property Prediction for Multicomponent Concentrated Aqueous Electrolyte Solutions. *Talanta* **81**, 149-155.

Rowland, D., May, P.M., 2012. An Investigation of Zdanovskii's Rule for Predicting the Water Activity of Multicomponent Aqueous Strong Electrolyte Solutions. *J. Chem. Eng. Data* **57**, 2589-2602.

Rowland, D., May, P.M., 2013. A Pitzer-based Characterization of Aqueous Magnesium Chloride, Calcium Chloride and Potassium Iodide Solution Densities to High Temperature and Pressure. *Fluid Phase Equilib.* **328**, 54-62.

Rowland, D., Königsberger, E., Hefter, G. and May, P.M. 2014. Aqueous Electrolyte Solution Modelling: Some limitations of the Pitzer equations. *Applied Geochemistry*, http://dx.doi.org/10.1016/j.apgeochem.2014.09.021.

Sillén, L.G., 1961. The Physical Chemistry of Seawater. in: Sears, M. (Ed.), Oceanography, Amer. Assoc. Adv. Sci. Publ. 67., Am. Assoc. Adv. Sci., Washington, DC, U.S.A., pp. 549-581.

Sillén, L.G., 1967a. How Have Sea Water and Air Got Their Present Compositions?. *Chem. Br.* **3**, 291.

Sillén, L.G., 1967b. The Ocean as a Chemical System. *Science* **156**, 1189-1197.

Turner, D.R., Whitfield, M., 1987. An Equilibrium Speciation Model for Copper in Sea and Estuarine Waters at 25 °C Including Complexation with Glycine, EDTA and NTA. *Geochim. Cosmochim. Acta* **51**, 3231-3239.

Turner, D.R., Whitfield, M., Dickson, A.G., 1981. The Equilibrium Speciation of Dissolved Components in Fresh Water and Seawater at 25 °C and 1 atm Pressure. *Geochim. Cosmochim. Acta* **45**, 855-881.

Weber, C.L., Vanbriesen, J.M., Small, M.S., 2006. A Stochastic Regression Approach to Analyzing Thermodynamic Uncertainty in Chemical Speciation Modeling. *Environ. Sci. Technol.* **40**, 3872-3878.

Whitfield, M., 1973. A Chemical Model for the Major Electrolyte Component of Seawater Based on the Brønsted-Guggenheim Hypothesis. *Mar. Chem.* **1**, 251-266.

Zittlau, A.H., Shi, Q., Boerio-Goates, J., Woodfield, B.F., Majzlan, J., 2013. Thermodynamics of the Basic Copper Sulfates Antlerite, Posnjakite, and Brochantite. *Chem. Erde* **73**, 39-50.

JOINT
EXPERT
SPECIATION
SYSTEM

Highlights

- Powerful chemical speciation code and database facility described
- Thermodynamic data harmonisation and automatic consistency-checking implemented
- Metal-ligand concentrations and solubilities in seawater and biofluids calculated
- Metastable equilibria included in aquatic chemistry modelling
- Limitations of ion association frameworks caused by specific-ion interactions